

Stochastic Optimal Control Models for Online Stores

Milan Bradonjić* and Albert Cohen†

January 14, 2013

Abstract

We present a model for the optimal design of an online auction/store by a seller. The framework we use is a stochastic optimal control problem. In our setting, the seller wishes to maximize her average wealth level, where she can control her price per unit via her reputation level. The corresponding Hamilton-Jacobi-Bellmann equation is analyzed for an introductory case, and *pulsing* advertising strategies are recovered for resource allocation.

Keywords: Stochastic optimal control models; Online stores; Hamilton-Jacobi-Bellmann equation.

1 Introduction

Auctions are a natural way to assign resources to selfish agents who compete for resources. These include diverse examples such as: wireless spectrum, arts, electronic equipment, songs, edges of a publicly known network [8, 27], or some other digital goods [15]. Please refer to the monograph by Milgrom [23] for an overview of auction theory. Online auctions have been established as electronic mechanisms in commercial transactions. Over a very short time, the Internet has become a formidable tool to connect buyers and sellers separated by large physical distances.

In this electronic arena, online sales have increased in orders of magnitude beyond imagination. Online auctions, as a tool for commercial transactions, have

*Milan Bradonjić is with the Theoretical Division, and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, e-mail: milan@lanl.gov. Corresponding author.

†Albert Cohen is with the Department of Mathematics and the Department of Statistics, Michigan State University, East Lansing, MI 48824, USA, e-mail: albert@math.msu.edu.

been established and used for maximization of auctioneer's profit [1, 6, 12, 18, 34, 19, 13, 7, 11]. On the other hand, the growth of the online commercial environment has come at the price of anonymous transactions where buyers have less recourse in protesting bad or unfair service. Online auctions are substantially different than traditional offline auctions [9, 33, 24, 21], and apropos that fact their bidders behave differently [4]. The presence of feedback forums, such as Amazon, eBay, and Yahoo!, at online auction sites, provide possibilities to rank other agents, get informed, and be ranked. Naturally, the online comment form has become vital in choosing which seller to purchase from. If the seller's reputation is damaged via negative feedback, she may expect a smaller sales rate. Reputation should simply mean that the higher the reputation of an agent is, the more confident that agent is. The notion of reputation helps agents, or the system designer, to optimally make their decisions, or to optimally design the mechanism, respectively [16, 10]. The notion that reputation affects the market outcome in the online settings has been studied [3, 31, 30], and there has been study on the optimal bidding strategy in sequential auctions [2].

In an online store, if the agent has profited greatly from dishonest transactions, or even decreased spending of resources on (a) advertising or (b) following up on sales to encourage higher reputation scores, this immediate increase in wealth may offset the long-term damage from lower reputation. With this in mind, one considers the existence of an optimal long term strategy. Using such a strategy, the seller can maximize her average expected wealth at a time of her choosing. Our work approximates the actions of a seller as being continuous. Such an approximation serves as an entry point and has the benefit of closed form solutions in some cases and the power of functional and numerical analysis in others. The results we present are encouraging and intuitively satisfying, and suggest more study in both discrete and continuous settings.

We note that our version of reputation evolution follows the standard Nerlove-Arrow construction [26], first postulated in 1962, and then extended to a stochastic setting by Sethi [32], and others. In Section 2, we propose an optimal finite horizon strategy, based on certain modeling assumptions for a seller to maximize her final average wealth. In this case, we model reputation, or goodwill, as a geometric Brownian motion. This process is of course familiar to those who work in quantitative finance, and the reader can see that it does indeed fall within a Nerlove-Arrow construction [26], if the adapted control, the advertising rate, is taken to be proportional to reputation level. The resulting control is of switching, or *pulsing*, type and is familiar to those who work in optimal advertising strategies. In Section 3, i.e., in our second model, we propose a finite horizon strategy that is based on the empirical work of Mink and Seifert [25], and we use for our reputation evolution the model proposed by Rao [29] and used by many others in optimal advertising mod-

els, such as Raman's recent work in Boundary Value Problems in the Stochastic Optimal Control of Advertising [28].

2 Implicit Resource Allocation Mechanism: Finite Time Horizon

For the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, define B as the standard Brownian motion that lives on this space. Take the current state (W, R, μ) as the Wealth, Reputation, and measure of resources spent on advertising *or* promotion, respectively. In this scheme, R is a positive number reflecting customer satisfaction. By choosing to shift up to 100% percent, $\epsilon \ll 1$, of her resources from promotion to processing and back, the seller can influence her wealth and reputation levels via μ . Positive μ corresponds to a promotional state; $\mu < 0$ corresponds to a processing state. The evolution of wealth and reputation is modeled over an interval $[t, T]$, where $0 \leq t \leq T$, by the two-state Markov process:

$$\begin{aligned} dW_t &= G(R_t, \mu_t) dt, \\ dR_t &= R_t (\mu_t dt + \sigma dB_t), \\ \mu \in \mathcal{A}_\epsilon &\equiv \left\{ \mu \in \mathfrak{F}^R \mid \mathbb{P} \left[\sup_{t \leq s \leq T} |\mu_s| \leq \epsilon \right] = 1 \right\}. \end{aligned}$$

The general model that we propose is to solve for $u(W, R, 0)$, where $\rho > 0$ is the discount rate and

$$\begin{aligned} u(W, R, 0) &:= \sup_{\mu \in \mathcal{A}_\epsilon} \mathbb{E} [W_T \mid W_0 = W, R_0 = R], \\ &= W + \sup_{\mu \in \mathcal{A}_\epsilon} \mathbb{E} \left[\int_0^T e^{-\rho s} G(R_s, \mu_s) ds \mid R_t = R \right], \\ u(W, R, t) &:= W + e^{\rho t} v(R, t), \\ v(R, t) &:= \sup_{\mu \in \mathcal{A}_\epsilon} \mathbb{E} \left[\int_t^T e^{-\rho s} G(R_s, \mu_s) ds \mid R_t = R \right]. \end{aligned}$$

This control problem corresponds to a store with (a) no salvage value upon closing or transfer, see [28], and (b) no switching costs proportional to μ^2 as we have the a prior bound $|\mu| \leq \epsilon \ll 1$. Our approximation is to assume a large inventory with an *almost* constant number of units sold per unit time. Under this assumption, we consider the growth rate to be the product of two factors: $G(R, \mu) = (1 - \mu)h(R)$. This product, revenue per unit time, is the *price per unit* $h(R)$ multiplied by a *resource rate factor* $r(\mu) = 1 - \mu$, the number of units sold

processed per unit time. Such a factor $r < 1$ ($\mu > 0$) represents the diversion of resources to advertising in return for a larger reputation score and larger potential sales price per unit in the future. If the seller ignores her advertising duties and diverts resources to processing, i.e. if $\mu < 0$, then $r > 1$ and we have an immediate revenue increase in return for lower reputation score and lower future sales price. It is *implicitly* assumed that this immediate loss (gain) can be attributed to spending (absorbing) an extra 100μ percent of time/resources on raising reputation scores (processing.) A balance is sought between these two competing factors h and r .

Finally, we model the evolution of the seller's reputation as a Geometric Brownian motion. This is done to reflect the asset-like nature of goodwill, and is only one of many approaches. In the next section, we propose a related control problem that uses the stochastic Nerlove-Arrow model [26]. This model has been used with success in the analysis of optimally planning advertising campaigns, and the resulting control is a *pulsing*, or switching, strategy that is similar to the one found using our Geometric Brownian dynamics for goodwill (reputation). Of course, the reader might notice that the Geometric Brownian motion model for the controlled reputation mechanism is also a Nerlove-Arrow type dynamic [26], if the control is taken to be proportional to the current level of goodwill, and if the stochasticity added also has a diffusion term $\sigma(R)$ that is linearly proportional to reputation R , instead of just constant. With these modeling constraints, the corresponding optimal control problem becomes

$$v(R, t) = \sup_{\mu \in \mathcal{A}_\epsilon} \mathbb{E} \left[\int_t^T e^{-\rho s} (1 - \mu_s) h(R_s) ds \mid R_t = R \right].$$

Then, the corresponding Hamilton-Jacobi-Bellmann (HJB) equation is

$$\begin{aligned} \max_{-\epsilon \leq \mu \leq \epsilon} \left[\frac{\partial v}{\partial t} + \mu R \frac{\partial v}{\partial R} + \frac{\sigma^2}{2} R^2 \frac{\partial^2 v}{\partial R^2} + e^{-\rho t} (1 - \mu) h(R) \right] &= 0, \\ v(R, T) &= 0. \end{aligned}$$

Basic algebra now leads to the following set of equations

$$\begin{aligned} e^{-\rho t} h(R) + \frac{\partial v}{\partial t} + \frac{\sigma^2}{2} R^2 \frac{\partial^2 v}{\partial R^2} + \epsilon \left| R \frac{\partial v}{\partial R} - e^{-\rho t} h(R) \right| &= 0, \\ v(R, T) &= 0. \end{aligned} \quad (1)$$

For general functions $h(\cdot)$, we are committed to the theory of partial differential equations. In the case where $h(\cdot)$ takes on a special form, we obtain simple, closed form results, via separation of variables. For example, we consider a power law model in the next section.

2.1 Power Law Model for the Growth Rate

Consider the growth rate $G(R, \mu) = (1 - \mu)R^\gamma$. This case corresponds to an object that has no inherent value if the seller's reputation is zero. As $h(0) = 0$, we expect that $\lim_{R \rightarrow 0} v(R, t) = 0$. We can in fact show this via the following lemma.

Lemma 1. *For $G(R, \mu) = (1 - \mu)R^\gamma$, we have $\lim_{R \rightarrow 0} v(R, t) = 0$.*

Proof. Since $\mu \leq \epsilon$, standard SDE theory [17] implies that for all $t \leq s \leq T$, we have $\mathbb{P}[R_s \leq \tilde{R}_s] = 1$, where $d\tilde{R}_s = \tilde{R}_s(\epsilon ds + \sigma dW_s)$ and $\tilde{R}_t = R$. Hence

$$\begin{aligned} 0 &\leq v(R, t) \leq \int_t^T e^{-\rho s} (1 + \epsilon) \mathbb{E} \left[\tilde{R}_s^\gamma \mid \tilde{R}_t = R \right] ds \\ &\leq (1 + \epsilon) e^{-\rho t} R^\gamma \int_t^T e^{(-\rho - \gamma(\epsilon - \frac{1}{2}\sigma^2) + \frac{1}{2}\sigma^2\gamma^2)(s-t)} ds \xrightarrow{R \rightarrow 0} 0, \end{aligned}$$

which proves the assertion of the lemma. This can also serve as an upper bound for $v(R, t)$. \square

As it is sometimes possible with optimal control problems in finance and operations research, we assume an Ansatz for the solution of (1). Then, separation of variables leads to this solution-control candidate pair of the form

$$\begin{aligned} v(R, t) &= e^{-\rho t} \psi_{\gamma, \rho}(t) R^\gamma \\ \mu_t &= \epsilon \cdot \text{sgn}(\gamma \psi_{\gamma, \rho}(t) - 1). \end{aligned}$$

We now prove the following theorem.

Theorem 2. *Let*

$$v(R, t) = \sup_{\mu \in \mathcal{A}_\epsilon} \mathbb{E} \left[\int_t^T e^{-\rho s} (1 - \mu_s) R_s^\gamma ds \mid R_t = R \right].$$

Then $v(R, t) = e^{-\rho t} \psi_{\gamma, \rho}(t) R^\gamma$, where

$$\begin{aligned} 1 - \rho \psi_{\gamma, \rho} + \psi'_{\gamma, \rho}(t) + \frac{\sigma^2}{2} \gamma(\gamma - 1) \psi_{\gamma, \rho}(t) + \epsilon |\gamma \psi_{\gamma, \rho}(t) - 1| &= 0, \\ \psi_{\gamma, \rho}(T) &= 0. \end{aligned} \quad (2)$$

Proof. For ease of computation, consider the case $\gamma = 1, \rho = 0$. (The case $\gamma < 1, \rho > 0$ follows a similar computation.) Then basic algebra leads to

$$\begin{aligned} R + \frac{\partial v}{\partial t} + \frac{\sigma^2}{2} R^2 \frac{\partial^2 v}{\partial R^2} + \epsilon R \left| \frac{\partial v}{\partial R} - 1 \right| &= 0, \\ v(R, T) &= 0. \end{aligned}$$

The corresponding control policy is $\mu_s = \epsilon \cdot \text{sgn} \left(\frac{\partial v}{\partial R}(R_s, s) - 1 \right)$. The ordinary differential equation (2) reduces to

$$\begin{aligned} 1 + \psi'_{1,0}(t) + \epsilon |\psi_{1,0}(t) - 1| &= 0, \\ \psi_{1,0}(T) &= 0. \end{aligned}$$

Although this is a nonlinear boundary value problem, it is still of first order, and separation of variables leads to $d\psi_{1,0}/dt = 1 + \epsilon |\psi_{1,0} - 1|$. So, for some constant C_1 , the following is satisfied

$$2\epsilon t + C_1 = \begin{cases} 2 \ln(1 + \epsilon(1 - \psi_{1,0})), & \text{if } \psi_{1,0} < 1 \\ -2 \ln(1 - \epsilon(1 - \psi_{1,0})), & \text{if } \psi_{1,0} > 1. \end{cases}$$

Using the boundary condition $\psi_{1,0}(T) = 0$, we obtain

$$\psi_{1,0}(t) = \begin{cases} \frac{1+\epsilon}{\epsilon} (1 - e^{-\epsilon(T-t)}), & \text{if } T - \frac{1}{\epsilon} \ln(1 + \epsilon) \leq t \leq T \\ \frac{e^{\epsilon(T-t)} - (1 - \epsilon^2)}{\epsilon(1 + \epsilon)}, & \text{if } 0 \leq t \leq T - \frac{1}{\epsilon} \ln(1 + \epsilon) \end{cases}$$

and $\mu_t \equiv \epsilon \cdot \text{sgn} \left(\frac{\partial v}{\partial R} - 1 \right) = \epsilon \cdot \text{sgn}(\psi_{1,0}(t) - 1)$ implies

$$\mu_t = \begin{cases} \epsilon, & \text{if } 0 \leq t < T - \frac{1}{\epsilon} \ln(1 + \epsilon) \\ -\epsilon, & \text{if } T - \frac{1}{\epsilon} \ln(1 + \epsilon) < t \leq T. \end{cases}$$

Notice that $\psi_{\gamma,\rho} \in C^1(0, \infty)$ and that $R^\gamma \in C^\infty(0, \infty)$. It is straightforward to see that $\psi_{\gamma,\rho}(t)R^\gamma$ is in the standard class \mathcal{D} for verification (see [14]), and is thus a solution to the optimal control problem. In the case $\gamma = 1$, if the time horizon T satisfies $\epsilon^{-1} \ln(1 + \epsilon) < T < 1 + \epsilon^{-1} \ln(1 + \epsilon)$, then there is indeed a switching time. Otherwise, the optimal policy is either $\mu \equiv \epsilon$ or $\mu \equiv -\epsilon$, depending on the size of the time horizon T . Again, a similar region for T is expected if $\gamma < 1$. \square

Interestingly, this control is independent of the reputation level R , and depends only on the time remaining. Also, the reader may notice a similarity of our approach to the iconic Merton Portfolio problem [22]. In our model, we multiply the power law growth rate by the processing rate dependent on the control, and we bound the size of our control. Still, it is not surprising that the solution is obtained by the same method, separation of variables, and has the same power law dependence on the underlying asset, in our case reputation. Unfortunately, the general $h(R)$ case does not express the same solution form.

Consider now

$$\begin{aligned} \mu_t &= \epsilon \mathbb{I}\left(0 \leq t \leq T - \frac{1}{\epsilon} \ln(1 + \epsilon)\right) - \epsilon \mathbb{I}\left(T - \frac{1}{\epsilon} \ln(1 + \epsilon) \leq t \leq T\right) \\ dR_t &= R_t (\mu_t dt + \sigma dB_t). \end{aligned}$$

For the sake of completion, we now verify for $T - \frac{1}{\epsilon} \ln(1 + \epsilon) \leq t \leq T$ that

$$\psi_1(t)R = \mathbb{E} \left[\int_t^T (1 - \mu_s) R_s \, ds \mid R_t = R \right].$$

In fact, direct substitution leads to

$$\begin{aligned} \mathbb{E} \left[\int_t^T (1 - \mu_s) R_s \mid R_t = R \, ds \right] &= \int_t^T \mathbb{E} [(1 - \mu_s) R_s \mid R_t = R] \, ds \\ &= \int_t^T \mathbb{E} [(1 - (-\epsilon)) R_s \mid R_t = R] \, ds \\ &= (1 + \epsilon) \int_0^{T-t} \mathbb{E} [R_s \mid R_0 = R] \, ds \\ &= (1 + \epsilon) \int_0^{T-t} R e^{-\epsilon s} \, ds \\ &= R \frac{1 + \epsilon}{\epsilon} (1 - e^{-\epsilon(T-t)}) \end{aligned}$$

which is the same as the corresponding quantity in (4). A similar computation can be made for $0 \leq t < T - \ln(1 + \epsilon)/\epsilon$ to verify that the solution of the partial differential equation coincides with the expectation of the stochastic integral. Also note that as $\epsilon \rightarrow 0$, we have $\lim_{\epsilon \rightarrow 0} v(R, t) = (T - t)R$. This is expected, as $\mu_t \rightarrow 0$ and so $\mathbb{E}[R_s \mid R_t = R] \rightarrow R$.

The study of *bang-bang* optimal control problems has a deep history, along with many approaches. A very useful and well studied path can be found in the work of [5], among others. This method solves for the sign of the gradient of value function, ∇v , using a Girsanov transformation [5], and the general solution follows. For a more detailed study consult [20].

Also, the reader may notice a similarity of our approach to the iconic Merton Portfolio problem [22]. In our model, we multiply the power law growth rate by the processing rate dependent on the control, and we bound the size of our control. Still, it should not be too surprising that the solution is obtained by the same method, separation of variables, and has the same power law dependence on the underlying asset, in our case reputation. Unfortunately, the general $h(R)$ case does not express the same solution form, whether we use a geometric Brownian motion or stochastic Nerlove-Arrow dynamics [26], see Section 3.

For sake of comparison, one could also analyze this problem as a single variable problem; the decision of when to make a single switch. For completeness, we present this in Appendix and show that solution via this method is the same as the HJB approach.

3 Explicit Resource Allocation Mechanism: Empirical Model of Mink and Seifert

In the previous model, we assumed an implicit mechanism for generosity. However, an *explicit* mechanism has been proposed in the novel work of [25]. There, the authors propose and empirically justify a growth rate of

$$h(R) = A + C \left(1 - \frac{1}{\ln(e + R)} \right), \quad (3)$$

where A relates to the inherent value of the object for sale and C is a parameter to be fitted. This is accomplished by obtaining data using an auction robot and then computing a single regression, which gives $C = 2.50$ in (3). It should be noted that in many economic papers on the effect of reputation on bidding and final sales price, the Mink-Seifert model is the first one we found to give an explicit relationship between reputation and price. The Mink-Seifert model also suggests a multiple regression formula where other factors, such as shipping costs and whether a “buy-it-now” price is offered, are considered as well. In that case, $C = 1.93$ and the authors in [25] comment this implies “... *the correlation between a seller’s revenues and her feedback score can be attributed to a large part to the fact that highly experienced sellers both have a higher feedback score and design the auction more favorably*”. In fact, they show that the coefficient attributed to shipping costs is larger than one, implying that customers put a high value on shipping when deciding on their bids, and that savvy agents take this into consideration. Finally, they posit that the horizon T does not affect the revenue stream as much as the shipping cost and reputation factors, and so we consider an arbitrary finite horizon model here.

As an initial approach, we consider only the effects of reputation, and leave the more general model with shipping costs for future work. Also, we consider the reputation mechanism first suggested by the work of Nerlove and Arrow [26] and later generalized to stochastic settings by, for instance, [32, 29, 28]. The stochastic wealth growth rate is given by

$$\begin{aligned} dW_t &= G(R_t, \mu_t) dt \\ dR_t &= (\mu_t - \kappa R_t) dt + \sigma dB_t \\ G(R, \mu) &= p(\mu) \left(A + \gamma \left(1 - \frac{1}{\ln(e + R)} \right) \right), \end{aligned}$$

where κ is a proportionality constant first proposed in [26], and γ represents the maximum premium earned over the inherent value A by the seller due to reputation. Given our initial approach in the previous section, where we took $p(\mu) = 1 - \mu$,

for $\mu \in [-\epsilon, \epsilon]$, we can now extend the processing rate $p(\mu)$ given expenditure. For a finite horizon, a company may be interested in capping its expenditure rate, and may only siphon a small amount away from the advertising budget as well in order to allocate that resource to processing instead. In general, we can have a rate $p(\mu)$ and a maximum possible processing rate M that satisfy

$$\begin{aligned}\lim_{\mu \rightarrow 0} p(\mu) &= 1, \\ \lim_{\mu \rightarrow +\infty} p(\mu) &= 0, \\ \lim_{\mu \rightarrow -\infty} p(\mu) &= M.\end{aligned}$$

Our previous control problem (2) returned a switching, or *pulsing*, type advertising campaign that was only time dependent. This made the subsequent analysis to verify the HJB solution easier. With the stochastic Nerlove-Arrow [26] dynamics for reputation and Mink-Seifert [25] growth rate for sales, we expect that switching would depend on both reputation and time. To investigate this, we again consider: (i) $p(\mu) = 1 - \mu$, and (ii) a seller who wishes to cap her advertising expenditure, but now as a total rate that is not necessarily proportional to her current reputation level

$$V(R, t) = \sup_{-\epsilon \leq \mu \leq \epsilon} \mathbb{E} \left[\int_t^{T \wedge \tau_0} (1 - \mu_s) \left(A + \gamma \left(1 - \frac{1}{\ln(e + R_s)} \right) \right) ds \mid R_t = R \right],$$

where $\tau_0 = \inf \{t \geq 0 \mid R_t = 0\}$.

The corresponding nonlinear HJB boundary value problem is

$$\begin{aligned}\max_{-\epsilon \leq \mu \leq \epsilon} \left[\frac{\partial V}{\partial t} + (\mu - \kappa R) \frac{\partial V}{\partial R} + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial R^2} + (1 - \mu) \left(A + \gamma \left(1 - \frac{1}{\ln(e + R)} \right) \right) \right] &= 0, \\ v(R, T) &= 0, \\ v(0, t) &= 0.\end{aligned}$$

The candidate for optimal control is again a pulsing strategy, switching between the values ϵ and $-\epsilon$ is given by

$$\mu = \epsilon \cdot \operatorname{sgn} \left(\frac{\partial v}{\partial R} - \left(A + \gamma \left(1 - \frac{1}{\ln(e + R)} \right) \right) \right).$$

The resulting boundary value problem becomes

$$\begin{aligned}
A + \gamma \left(1 - \frac{1}{\ln(e + R)} \right) + \frac{\partial V}{\partial t} - \kappa R \frac{\partial V}{\partial R} + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial R^2} \\
+ \epsilon \left| \frac{\partial v}{\partial R} - \left(A + \gamma \left(1 - \frac{1}{\ln(e + R)} \right) \right) \right| &= 0, \\
v(R, T) &= 0, \\
v(0, t) &= 0.
\end{aligned}$$

Naturally, the strategy will switch dependent on both the time and current reputation state. This is different from the power law model as that led a time-only dependent pulsing strategy. Of course, the solution of the HJB equation is only a candidate for the solution of the control problem, and it must be verified as in the previous section. The solution of this HJB equation, however, may very well be non-smooth, and so the notion of generalized solutions (see [14]) must be visited. Further analysis on this equation is forthcoming.

4 Conclusion and Future Work

In this work, we defined a general framework for the problem of selling goods online when buyer feedback factors into the sales rate. The framework was a stochastic optimal control problem, where the seller wishes to maximize her average wealth level at a fixed time of her choice. We presented a method to incorporate the optimal design of an online auction (online store) by a seller in the presence of reputation management. Then, we introduced and analyzed the corresponding Hamilton-Jacobi-Bellmann equation. To obtain intuition, we first analyzed a model where the revenue per sale, or the price per unit, was dependent solely on reputation, and multiplied by a mark-down factor. Such a factor represented the loss per item the seller could expect for behaving generously today in return for a larger reputation score and larger potential revenue in the future. As an example, we considered a power law growth rate function, which again provided insight into the general solution method.

Subsequently, we have shifted our attention to an empirically justified model, and proposed that a seller might optimally design her online store in accordance with the Mink-Seifert model, an *explicit* mechanism proposed in the novel work of [25]. There, the authors propose and empirically justify a price per unit of $h(R) = A + C(1 - 1/\ln(e + R))$, where A relates to the inherent value of the object for sale and C is a parameter to be fitted that represents the maximum premium expected due to reputation over inherent value. This is accomplished by obtaining data using

an auction robot and then computing a single regression, which gives $C = 2.50$. It should be noted that in many economic papers on the effect of reputation on bidding and final sales price, the Mink-Seifert model is the first one we found to give an explicit relationship between reputation and price.

We have incorporated the reputation mechanism first suggested by the work of Nerlove and Arrow [26] and later generalized to stochastic settings in [32, 29, 28]. With these modeling considerations, the bivariate Revenue-Reputation Markov process. Naturally, the strategy is expected to switch dependent on both the time and the current reputation state. This is different from the power law model, which led to a time-only dependent pulsing strategy. Further analysis on the resulting HJB, including numerics, are forthcoming.

Acknowledgments. The authors thank Nir Gavish and Steven Shreve for useful discussions. Milan Bradonjić is supported by the Department of Energy ASCR program at Los Alamos National Laboratory under contract DE-AC52-06NA25396.

References

- [1] AGGARWAL, G., GOEL, A., AND MOTWANI, R. Truthful auctions for pricing search keywords. In *Proc. of EC '06* (2006), pp. 1–7.
- [2] ARORA, A., XU, H., PADMAN, R., AND VOGT, W. Optimal bidding in sequential online auctions. Working Paper 2003-4. Carnegie Mellon University.
- [3] BAJARI, P., AND HORTACSU, A. Economic insights from internet auctions: A survey. NBER Working Papers 10076, National Bureau of Economic Research, Inc, Nov. 2003.
- [4] BEAM, C., AND SEGEV, A. Auctions on the internet: A field study, 1998.
- [5] BENEŠ, V. E. Composition and invariance methods for solving some stochastic control problems. *Advances in Applied Probability* 7 (1975), 299–329.
- [6] BÖRGERS, T., COX, I., PESENDORFER, M., AND PETRICEK, V. Equilibrium bids in sponsored search auctions: Theory and evidence. *Technical report, University of Michigan* (2007).
- [7] CARY, M., DAS, A., EDELMAN, B., GIOTIS, I., HEIMERL, K., KARLIN, A. R., MATHIEU, C., AND SCHWARZ, M. Greedy bidding strategies for keyword auctions. In *EC '07: Proceedings of the 8th ACM conference on Electronic commerce* (New York, NY, USA, 2007), ACM, pp. 262–271.

- [8] CARY, M. C., FLAXMAN, A. D., HARTLINE, J. D., AND KARLIN, A. R. Auctions for structured procurement. In *SODA '08: Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms* (Philadelphia, PA, USA, 2008), Society for Industrial and Applied Mathematics, pp. 304–313.
- [9] CASSADY, R. J. *Auctions and Auctioneering*. University of California Press, Berkeley and Los Angeles, 1967.
- [10] DELLAROCAS, C. Reputation mechanism design in online trading environments with pure moral hazard. *Info. Sys. Research* 16, 2 (2005), 209–230.
- [11] EDELMAN, B., AND OSTROVSKY, M. Strategic bidder behavior in sponsored search auctions. *Decis. Support Syst.* 43, 1 (2007), 192–198.
- [12] EDELMAN, B., OSTROVSKY, M., AND SCHWARZ, M. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review* 97, 1 (March 2007), 242–259.
- [13] FENG, J., BHARGAVA, H. K., AND PENNOCK, D. M. Implementing sponsored search in web search engines: Computational evaluation of alternative mechanisms. *Inform. Journal on Computing* (2006).
- [14] FLEMING, W. H., AND RISHEL, R. W. *Deterministic and stochastic optimal control*. Springer-Verlag, Berlin and New York, 1975.
- [15] GOLDBERG, A. V., HARTLINE, J. D., AND WRIGHT, A. Competitive auctions and digital goods. In *In Proceedings of the 12th Symposium on Discrete Algorithms* (2001), ACM/SIAM, pp. 735–744.
- [16] HOUSER, D., AND WOODERS, J. Reputation in auctions: Theory, and evidence from ebay. *Journal of Economics & Management Strategy, Blackwell Publishing*, vol 15, 2 (2006), 353–369.
- [17] KARATZAS, I., AND SHREVE, S. E. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, 1988.
- [18] LAHAIE, S. An analysis of alternative slot auction designs for sponsored search. In *Proc. of EC '06* (New York, NY, USA, 2006), ACM Press, pp. 218–227.
- [19] LAHAIE, S., AND PENNOCK, D. M. Revenue analysis of a family of ranking rules for keyword auctions. In *Proc. of EC '07* (New York, NY, USA, 2007), ACM Press, pp. 50–56.

- [20] LIBERZON, D. *Switching in systems and control, Systems and Control: Foundations and Applications*. Birkhuser Boston Inc., 2003.
- [21] MCAFEE, R. P., AND MCMILLAN, J. Auctions and bidding. *Journal of Economic Literature* 25 (1987), 699–738.
- [22] MERTON, R. C. Lifetime portfolio selection under uncertainty: The continuous-time case. *Rev. Economics and Statistics* 51 (1969), 247–257.
- [23] MILGROM, P. *Putting Auction Theory to Work*. Cambridge University Press, Cambridge, 2004.
- [24] MILGROM, P., AND WEBER, R. A theory of auctions and competitive bidding. *Econometrica* 50 (1982), 1089–1122.
- [25] MINK, M., AND SEIFERT, S. Reputation on ebay and its impact on sales prices. In *Group Decision and Negotiation (GDN) (International Conference Karlsruhe, Germany, June 25-28, 2006)* (2006), pp. 253–255.
- [26] NERLOVE, M., AND ARROW, K. J. Optimal advertising policy under dynamic conditions. *Economica* 29 (1962), 129–142.
- [27] NISAN, N., AND RONEN, A. Algorithmic mechanism design. *Games and Economic Behavior* 35 (2001), 166–196.
- [28] RAMAN, K. Boundary value problems in stochastic optimal control of advertising. *Automatica* 42 (August 2006), 1357–1362.
- [29] RAO, R. C. Estimating continuous time advertising-sales models. *Marketing Science* 5, 2 (1986), 125–142.
- [30] RESNICK, P., AND ZECKHAUSER, R. Trust among strangers in internet transactions: Empirical analysis of ebay’s reputation system. In *The Economics of the Internet and E-Commerce*, M. R. Baye, Ed. Volume 11 of *Advances in Applied Microeconomics*. Amsterdam, Elsevier Science, 2002, pp. 127–157.
- [31] RESNICK, P., ZECKHAUSER, R., SWANSON, J., AND LOCKWOOD, K. The value of reputation on ebay: A controlled experiment. *Experimental Economics* 9, 2 (2006), 79–101.
- [32] SETHI, S. P. Deterministic and stochastic optimization of a dynamic advertising model. *Optimal Control Applications and Methods* 4 (1983), 179–184.

- [33] STARK, R. M., AND ROTHKOPF, M. H. Competitive bidding: A comprehensive bibliography. *Operations Research* 27 (1979), 364–390.
- [34] VARIAN, H. R. Position auctions. *International Journal of Industrial Organization* 25, 6 (2007), 1163–1178.

Appendix

Here we consider the 1-switching time problem. More precisely, consider $dR_t = R_t(\mu_t dt + \sigma dB_t)$, with known R_s and constant μ . Then it easily follows

$$\mathbb{E}[R_t \mid R_s = R] = Re^{\mu(t-s)}. \quad (4)$$

Now, consider

$$\mu_\tau = \begin{cases} \epsilon, & 0 \leq \tau \leq t \\ -\epsilon, & t < \tau \leq 1. \end{cases}$$

Hence,

$$\begin{aligned} f(R, t) &\equiv \mathbb{E}\left[\int_0^1 (1 - \mu_s)R_s ds \mid R_0 = R\right] \\ &= \mathbb{E}\left[\int_0^t (1 - \epsilon)R_s ds \mid R_0 = R\right] + \mathbb{E}\left[\int_t^1 (1 + \epsilon)R_s ds \mid R_0 = R\right] \\ &= (1 - \epsilon)\mathbb{E}\left[\int_0^t R_s ds \mid R_0 = R\right] + (1 + \epsilon)\mathbb{E}\left[\int_t^1 R_s ds \mid R_0 = R\right] \\ &= (1 - \epsilon)\int_0^t \mathbb{E}[R_s \mid R_0 = R] ds + (1 + \epsilon)\int_t^1 \mathbb{E}[R_s \mid R_0 = R] ds \quad (5) \end{aligned}$$

where the last equality is obtained by the tower property of the conditional expectations. Furthermore, by (4), the expression (5) equals:

$$\begin{aligned} f(R, t) &= (1 - \epsilon)\int_0^t Re^{\epsilon s} ds + (1 + \epsilon)\int_t^1 \mathbb{E}[R_t e^{-\epsilon(s-t)} \mid R_0 = R] ds \\ &= (1 - \epsilon)R\frac{e^{\epsilon t} - 1}{\epsilon} + (1 + \epsilon)\int_t^1 e^{-\epsilon(s-t)} Re^{\epsilon t} ds \\ &= R\left(\frac{1 - \epsilon}{\epsilon}(e^{\epsilon t} - 1) + \frac{1 + \epsilon}{\epsilon}(e^{\epsilon t} - e^{2\epsilon t - \epsilon})\right). \end{aligned}$$

For $f(R, t)$ defined in (5), we now seek for the absolute maximum on $[0, 1]$. Equating its first derivative with zero, $f_t(R, t_1) = 0$, gives $t_1 = 1 - \ln(1 + \epsilon)/\epsilon$ and

$f(R, t_1) = R(e^\epsilon + \epsilon^2 - 1)/(\epsilon(1 + \epsilon))$. Collecting these into one computation, we obtain

$$\begin{aligned} f_t(R, 0) &= R \frac{(1 + \epsilon)(1 - e^{-\epsilon})}{\epsilon} \\ f_t(R, t_1) &= R \frac{e^\epsilon + \epsilon^2 - 1}{\epsilon(1 + \epsilon)} \\ f_t(R, 1) &= R \frac{(1 - \epsilon)(e^\epsilon - 1)}{\epsilon}. \end{aligned}$$

Now, simple analysis shows that $f(R, 1) < f(R, 0) < f(R, t_1)$, for all $\epsilon \in (0, 1]$. This is, of course, the result obtained in the HJB analysis in Section 2.1 for $h(R) = R$ and $T = 1$. In other words, for $h(R) = R$,

$$v(R, 0) = \sup_{0 \leq s \leq 1} f(R, s).$$